Position- and Base Pair-specific Comparison of *p53* Mutation Spectra in Human Tumors: Elucidation of Relationships between Organs for Cancer Etiology

Werner K. Lutz, Tamas Fekete, and Spiros Vamvakas

Department of Toxicology, University of Würzburg, Würzburg, Germany

A new approach to analyze the *p53* mutation database of the European Molecular Biology Laboratory for a comparison of mutation spectra is described, with the aim of investigating organ specificity of etiological factors and putative organ-to-organ relationships in cancer pathogenesis. The number of entries of each nucleotide- and base-pair substitution-specific mutation was divided by the total number of tumors analyzed. For each organ pair, the difference of the frequencies was calculated (absolute values) and, going through all mutations, the sum of the mutation-specific frequency differences was calculated. Resulting values could range from 0 (full concordance) to 2 (full discordance). Skin, lung, and urinary bladder showed highly independent mutation spectra (maximum discordance value = 1.48 for skin versus brain), in agreement with the presence of specific factors responsible for a large number of the respective tumors (UV light, smoking, aromatic amines). The three organs with the smallest sum of discordance values were mammary gland (breast), colon, and esophagus. The minimum organ-to-organ discordance value was 0.95, for stomach versus colon. For these organs, common, possibly also endogenous, cancer risk factors could be postulated as contributing to the observed mutation spectrum. The remaining cancers (ovary, sarcoma, leukemia/lymphoma, brain, head and neck, and stomach, in order of increasing discordance) were of intermediate range and showed a mix of values. Reasons for close relationship to some of the other organs and marked differences to others are discussed. Exclusion of the "hot-spot" mutations did not markedly alter the observed relationships, indicating that a putative selective growth advantage does not cover up the etiological basis for the observed mutation spectrum. It is expected that much more insight into carcinogenesis and cancer could be gained by further exploratory analyses of mutation databases. *Key words*: cancer, cancer risk factor, mutation spectrum, organ specificity, *p53*, tumor etiology, tumor-suppressor gene. *Environ Health Perspect* 106:207–211 (1998). [Online 10 March 1998]
http://ehpnet1.niehs.nih.gov/docs/1998/106p207-211lutz/abstract.html

The *p53* tumor-suppressor gene is mutated in approximately 50% of human cancers, and the spectra of substitution mutations have been analyzed to elucidate tumor pathogenesis (*1,2*). The major issue has been the identification and discussion of the most frequently mutated codon sites ("hot spots") and the analysis of the prevalence of the various types of base-pair substitutions.

A database of all published *p53* substitution mutations in human tumors and cell lines is available on-line through the European Bioinformatics Institute or through the International Agency for Research on Cancer (*3*). More than 5,000 mutations had been collected by 1996. The database represents a unique source of information that can be analyzed in many different ways. A recent example concerning the organ specificity has been published by Lasky and Silbergeld in *EHP* (*4*).

Base-pair substitutions are the result of mispairing during DNA replication. This mispairing can be due to primary DNA lesions, such as DNA adducts, or to spontaneous replication errors. DNA adducts are formed both by exogenous carcinogens (*5*) and by endogenous agents (*6*). The adduct profile is not only

carcinogen and base specific but appears to be position specific also. For instance, site-targeted formation of carcinogen–DNA adducts contributes to the *p53* mutation spectrum seen in lung cancer (*7*). Therefore, if mutation spectra can be analyzed not only for the frequency of specific types of base-pair substitution but also in terms of position, this might improve the analysis of relationships between organs with respect to the etiology of the respective tumors.

Methods

Database and selection of entries. The database was accessed on 17 January 1997 at *http://www.eb.ac.uk*, and 5,174 entries were retrieved. They were converted to a Microsoft Excel file, which formed the basis of all selection procedures and calculations described. Lines that contained the following names of organs or cancers in column G, "tumor type," were used (number of recovered entries in parentheses): urinary bladder (178), brain (207), breast (331), colon (461), esophagus (146), head and neck (244), leukemia/lymphoma (349), lung (453), ovary (218), sarcoma (110), skin (125), and stomach (194). Misspelled entries were not

considered. The database was ordered by the nucleotide position (145 to 1024; i.e., from the first nucleotide of codon 49 to the first nucleotide of codon 342) and, for each nucleotide, by the specific base change noted in column E (for instance, 469 G to A, 469 G to C, 469 G to T).

Analysis of the relationship between two organs. For each of the 12 organs (= 12 columns), we divided the number of entries for each mutation by the total number of tumors analyzed. These frequencies indicated the relative abundance of a specific mutation. For all organ pairs, we calculated the (absolute) difference of the relative frequencies, and, going through all mutations, calculated the sum of the differences. The procedure is illustrated in Table 1 for nucleotide 469 (i.e., for the first base of codon 157). For the comparison between colon and esophagus, the G to T transversion of nucleotide 469 has practically the same frequency in the two organs, and the difference is minimal ($|0.0065 - 0.0068| = 0.0003$). On the other hand, the same mutation displays a fivefold higher relative frequency in lung tumors as compared to colon tumors ($|0.0375 - 0.0065| = 0.0310$). Large differences in many positions and substitutions result in large discordance values when the sum of the differences is calculated. In the case of full concordance of the mutational spectrum between two organs, the sum of the differences is 0. For complete discordance, a value of 2 will result: if a relative frequency >0 in one organ is always opposite a relative frequency of 0 in the other organ, the sum of the differences will be (1+1 =) 2.

Organs considered. A preliminary analysis which included the entries for cervix (cer; 31 entries), pancreas (pan; 94), prostate (pro; 46), renal (kid; 36), thyroid (thy; 32), and uterus (ute; 36) revealed that organs with only a small number of entries showed larger sums of discordance values than organs with a large number of analyzed tumors (Fig. 1). This might be due to the fact that with a small sample, the minimum relative frequency is

larger than with a large sample. We limited the analysis to those organs for which at least 100 mutations had been recorded. In this group, only the skin could have been affected by the artifact. However, another explanation for the large discordance value of skin is put forward below; its inclusion in the analysis was considered appropriate.

Results

The organ-to-organ relationship for the mutation spectra in the *p53* gene was analyzed by calculating the relative frequencies of position- and base pair-specific substitution mutations for each organ, and then summing up organ-to-organ differences of the frequencies over all mutations. The results are shown in Table 2 in the form of a 12 × 12 matrix containing a set of 12 × 11 = 132 organ-to-organ values in duplicate. Small values indicate small differences between the mutation spectra (i.e., concordance); large values indicate larger differences (i.e., discordance). The theoretical span of

values is from 0 (full concordance) to 2 (complete discordance). Observed values ranged from a minimum of 0.95 (colon compared with stomach) to a maximum of 1.48 (skin compared with brain). The frequency distribution of the values is shown in Figure 2. It is defined by a mean and standard deviation of 1.23 ± 0.14. The median is 1.26, the mode is 1.30, the lower quartile (25th percentile) is at 1.10, and the upper quartile (75th percentile) is at 1.36. The distribution therefore is quite symmetrical.

For the following discussion, the values shown in Table 2 have been grouped. The first and fourth quartiles of the values are marked; the two intermediate quartiles are not marked specifically. A number of observations can be made on this basis. The three organs that exhibit the largest sum of discordance values are the skin, lung, and urinary bladder, and the three organs with the smallest sum of values are the mammary gland, colon, and esophagus. The remaining organs/cancers (sarcoma, stomach, head and

neck, brain, leukemia/lymphoma, ovary, in the order of decreasing discordance) range in between and show a mix of values, with close relationship to some of the other organs and marked differences to others.

The first group can be characterized by well-known, specific high-risk factors: sunlight for skin, cigarette smoking for lung, and aromatic amines for the urinary bladder. Those at the other end of the ranking are in part hormone dependent (mammary, ovary) or have multiple risk factors (esophagus, colon). Most of these also appear to have an endogenous component (8). This analysis therefore could provide some insight into the question of the relative contributions of different factors to carcinogenesis.

Specific Etiology

Values in the fourth quartile were obtained most frequently for skin tumors (against all other organs). Exposure to sunlight is the

Table 1. Procedure for the analysis of the organ-to-organ relationships for *p53* mutation spectra in tumors, exemplified for the first nucleotide of codon 157, nucleotide 469 (G)

|	Base change		
	G→A	G→C	G→T
No. of tumors with mutation			
Esophagus (146 tumors analyzed)	1	0	1
Colon (461 tumors analyzed)	0	2	3
Lung (453 tumors analyzed)	0	0	17
Frequencies of tumors with mutation			
Esophagus	0.0068 (1/146)	0	0.0068
Colon	0	0.0043	0.0065
Lung	0	0	0.0375
Organ-to-organ difference of frequencies (absolute values)			
Esophagus to colon	0.0068	0.0043	0.0003
Sum for nucleotide 469	0.0114		
Sum for all nucleotides	1.03[a]		
Colon-to-lung	0	0.0043	0.0310
Sum for nucleotide 469	0.0353		
Sum for all nucleotides	1.37[a]		

[a]Value for Table 2.
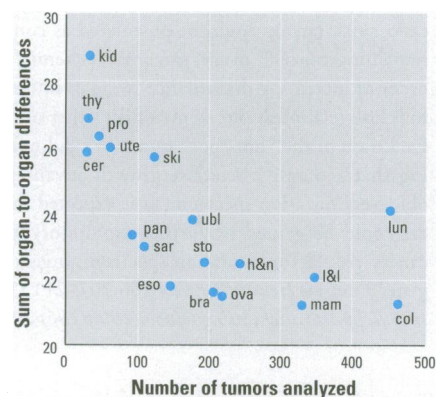
Figure 1. Sum of organ-to-organ discordance values for 18 organs, as a function of the number of tumors analyzed. Abbreviations: ubl, urinary bladder; bra, brain; mam, mammary gland (breast); col, colon; eso, esophagus; h&n, head and neck; l&l, leukemia and lymphoma; lun, lung; ova, ovary; sar, sarcoma; ski, skin; sto, stomach; kid, renal; thy, thyroid; pro, prostate; ute, uterus; cer, cervix; pan, pancreas.

Table 2. Discordance values for the relationships between organs with respect to the *p53* mutation spectra in human tumors (the procedure for the calculation of the values is exemplified in Table 1)

	ubl	bra	mam	col	eso	h&n	l&l	lun	ova	sar	ski	sto
ubl	0	1.33	1.18	1.30	1.25	1.31	1.28	1.32	1.33	1.30	1.44[a]	1.41[a]
bra	1.33	0	1.12	1.00[b]	1.18	1.26	1.09[b]	1.39[a]	1.06[b]	1.12	1.48[a]	1.09[b]
mam	1.18	1.12	0	1.06[b]	1.06[b]	1.03[b]	1.05[b]	1.13	1.00[b]	1.18	1.40[a]	1.19
col	1.30	1.00[b]	1.06[b]	0	1.03[b]	1.26	1.01[b]	1.37[a]	1.07[b]	1.26	1.40[a]	0.95[b]
eso	1.25	1.18	1.06[b]	1.03[b]	0	1.04[b]	1.10[b]	1.19	1.11	1.20	1.41[a]	1.19
h&n	1.31	1.26	1.03[b]	1.26	1.04[b]	0	1.16	1.15	1.20	1.24	1.37[a]	1.30
l&l	1.28	1.09[b]	1.05[b]	1.01[b]	1.10[b]	1.16	0	1.31	1.12	1.28	1.45[a]	1.23
lun	1.32	1.39[a]	1.13	1.37[a]	1.19	1.15	1.31	0	1.21	1.37[a]	1.42[a]	1.43[a]
ova	1.33	1.06[b]	1.00[b]	1.07[b]	1.11	1.20	1.12	1.21	0	1.24	1.46[a]	1.08[b]
sar	1.30	1.12	1.18	1.26	1.20	1.24	1.28	1.37[a]	1.24	0	1.43[a]	1.30
ski	1.44[a]	1.48[a]	1.40[a]	1.40[a]	1.41[a]	1.37[a]	1.45[a]	1.42[a]	1.46[a]	1.43[a]	0	1.39[a]
sto	1.41[a]	1.09[b]	1.19	0.95[b]	1.19	1.30	1.23	1.43[a]	1.08[b]	1.30	1.39[a]	0
Sum	14.44[a]	13.11	12.39[b]	12.71[b]	12.77[b]	13.31	13.09	14.28[a]	12.87	13.92	15.66[a]	13.56

Abbreviations: ubl, urinary bladder; bra, brain; mam, mammary gland (breast); col, colon; eso, esophagus; h&n, head and neck; l&l, leukemia and lymphoma; lun, lung; ova, ovary; sar, sarcoma; ski, skin; sto, stomach.
[a]Values belonging to the fourth quartile (>75th percentile).
[b]Values belonging to the first quartile (<25th percentile).

most important exogenous risk factor for skin tumor induction. This factor does not operate in the other tissues examined here. Thus, the present analysis allowed the identification of an organ where tumor formation is based primarily on risk factors that do not act in other organs. It also indicates that mutation spectra in the tumor retain etiological information, although the phenotypic selection of mutations is expected to bring about a modification of the carcinogen-derived spectrum in the normal tissue. The fact that the mutation spectra of skin tumors are clearly different from all other investigated human cancers but similar to those observed in UV-light model systems was also demonstrated in a recent comparison of *p53* mutations in skin cancers with those in all kinds of internal cancers (*9,10*).

The lung exhibited the second most large organ-to-organ discordance values. Tobacco smoking is the prominent organ-specific risk factor. However, smoking is also involved in the formation of other tumors, such as tumors of the head and neck and the esophagus. Why were these organs not closely related to the lung in the present analysis? One explanation might lie in the fact that it is not smoking alone, but the combination of smoking and excessive alcohol drinking which is responsible for some 75% of esophageal and 90% of head and neck cancers, whereas alcohol does not appear to be a risk factor in lung cancer (*11,12*). Furthermore, the result of the present analysis could support the idea that the smoke-related carcinogens which dominate the process in the lung are different from those operating in pharynx, larynx, and esophagus.

Tumors of the urinary bladder also appeared to have little in common with other organs, including the lung (organ-to-organ value 1.32), although cigarette smoking is also a risk factor for bladder cancer. However, the association with smoking is much stronger for lung cancer than for bladder cancer. The relative risk for smokers compared to nonsmokers to develop lung cancer is 11.3, whereas for bladder cancer the relative risk is only 2.2 (*13,14*). Second, as previously suggested by DNA adduct studies, aromatic amines are a major risk factor for bladder cancer (*15*), whereas carcinogenicity of cigarette smoking for the lung appears to be based primarily on other carcinogens, such as polynuclear aromatic hydrocarbons and nitroso compounds, and is strongly modulated by tumor-promoting aspects of smoking-associated toxicity (*16*). Hence, the dissimilarity in *p53* mutation spectra between the lung and the urinary bladder reflects both qualitative and quantitative differences in the role of cigarette smoke carcinogens for the two organs. This distinct *p53* mutation spectrum in bladder tumors was also reported in a previous comparison with the mutation spectra of smoking-related lung tumors and smoking-independent colon cancer (*17*).

Sarcomas also exhibited relatively little concordance with other cancers. The closest was with brain tumors with a value of 1.12. We cannot offer a sound hypothesis for any specific etiology of human sarcoma formation as indicated by the present analysis, especially in view of the fact that sarcomas arise in all kinds of tissues and exhibit different differentiation grades. The artifact of apparently large discordance values associated with low numbers of analyzed tumors may apply here. Sarcoma was the tumor class with the lowest number of entries (110) considered in this analysis. When more data

become available, this artifact should disappear, and it might be possible to stratify the entries by type and location to see whether relationships are hidden by the diversity of the tumors registered as sarcomas.

Related Organs

The small values in Table 2 indicate a relationship between the two organs considered, based on similar tumor etiology or related tumor biology. The lowest value was obtained for the pair colon and stomach. Numerous risk factors are discussed for the two sites. A high-fat/low-fiber diet appears to promote colon carcinogenesis, whereas gastritis, predominantly in connection with infection, is associated with tumorigenesis in the stomach. In both cases, exogenous DNA-reactive carcinogens cannot fully explain the process of carcinogenesis, and background DNA damage, together with tumor-promoting aspects such as the stimulation of cell division, might be more important in both tissues. Background DNA damage could be postulated, for instance, from oxidative stress and lipid peroxidation resulting in the increased formation of oxygen radicals and secondary products such as aldehydes. Such a common background mechanism may explain the close relationship between these two organs in terms of mutation spectra.

The colon showed the largest number of close relationships with, for example, the hormone-dependent tissues of the mammary gland and the ovary, as well as with brain and leukemia/lymphoma. More refined analyses will be required to investigate which of these relations are based on similar pathogenesis.

Tumors of the head and neck and the esophagus were also closely interrelated. This is in line with regular alcohol consumption, particularly when combined with cigarette smoking. Genotoxic aldehydes, both present directly in cigarette smoke as well as in intermediate metabolites of alcohol, could form the common link. Because aldehydes are also important secondary products of lipid peroxidation, this class of carcinogens could also account at least in part for other relationships, for which no sound explanation can be given at this time.

Leukemias/lymphomas displayed similarities with cancers of brain and colon, and large discordance to cancers having known exogenous risk factors, such as skin, lung, and bladder. C-G to T-A transitions at CpG sites, possibly as a result of oxidative deamination of 5-methylcytosine, were particularly frequent in tumors of the brain, colon, stomach, and endometrium and in hematological cancers (*1*). Hence, although specific chemicals and viruses may be contributing factors,
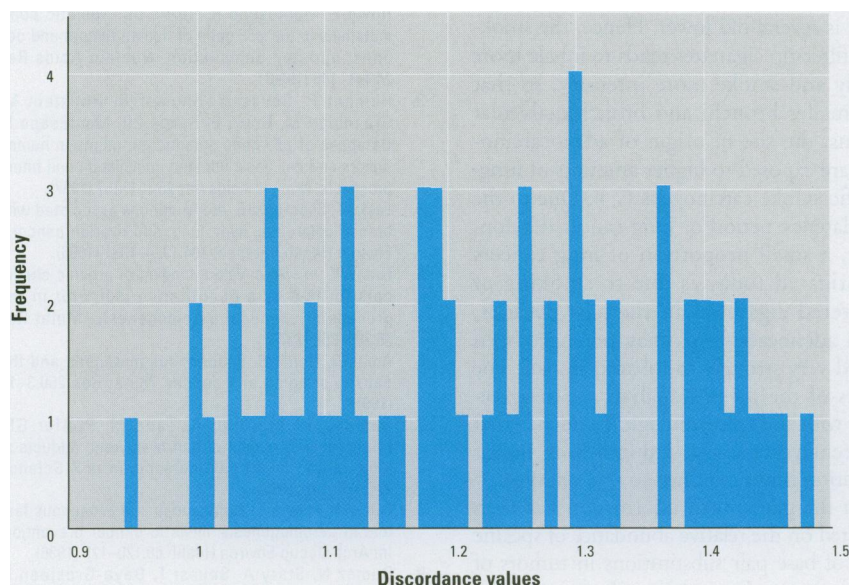

Figure 2. Frequency distribution of the organ-to-organ discordance values shown in Table 2.

the present analysis would support the hypothesis of a strong endogenous component in the etiology of leukemias and lymphomas.

The mutation pattern of tumors of the mammary gland (breast cancer) was clearly distinct only from the one in skin tumors (1.40) but revealed relationships above average with all other organs (<1.19) and even showed a value of 1.13 with the lung. The closest relationship was observed with the ovary, another hormone-dependent organ. Although exogenous factors contribute at least to a certain extent to the breast cancer risk, established factors cannot account for the majority of breast cancers. The association with high intake of animal fat is increasingly contested, and exposure to polycyclic aromatic hydrocarbons (PAHs) known to cause mammary gland cancer in animals has not been strongly associated with breast cancer in humans (18,19). The present analysis would support the multifactorial etiology without prominent specific factors.

Discussion

The database on p53 mutations in human tumors provides a unique opportunity to investigate the molecular epidemiology of cancer. In this paper, the question of whether the mutation spectra in different organs can be analyzed to reveal etiological similarities and differences was addressed.

The analysis identified organs with little relationship to others, reflecting a rather independent pathogenetic mechanism based on specific exogenous factors (skin, lung, bladder). On the other hand, the hormone-dependent tissues of the breast and ovary, gastrointestinal tumors, as well as leukemia/lymphoma and brain tumors showed substantial overlap. The hypothesis of a common pathway in the pathogenesis of these tumors, for instance by background DNA damage, could be formulated. This is supported by a descriptive epidemiological analysis that revealed relatively small differences between incidence rates reported worldwide for some of these cancers (8). However, it cannot be excluded that two organ-cancers have quite diverse etiology and still have a similar mutation spectrum. More information on the specificity of mutational fingerprints of different carcinogens will be required to shed more light on this question.

The mutation spectrum in a tumor is the result of two main factors: the nature of the DNA damage and the ability of the mutations formed to convey an advantage to the cell in terms of clonal expansion and tumor progression. The latter property results in so-called hot spots. The question

therefore is whether hot spots should be included in the present analysis of etiologic factors or whether they distort the situation and are better omitted.

In a recent analysis, hot spots were shown not to be universal for all organs, and the distribution of base-pair changes over hot spots also varied by cancer (4). Based on our analysis, for instance, the share among all mutations of the hot spot mutation CGT to TGT in codon 273 ranged from 0.4% in head and neck tumors to 10.6% in the brain. The CGT to CAT mutation in codon 273 contributed between 1.6% in lung and 6.7% in the mammary gland. Therefore, hot spots do retain some organ-specific features and could be included in the present type of analysis. In fact, when we restricted our analysis to those mutations that contributed less than 3% of all documented mutations, no additional information could be derived in terms of etiological specificity (data not shown). The inclusion of the hot-spot mutations in the present analysis therefore did not interfere with our interpretation. However, this is not a general rule and does not exclude the possibility that certain other questions may be better addressed by a separate analysis of hotspot and non-hot-spot-mutations.

Recent epidemiological studies with cigarette smokers indicated a markedly steeper rise of the incidence of adenocarcinoma compared with squamous cell carcinoma of the lung (20–22). In the past, the association between adenocarcinoma and cigarette smoking was considered to be weak or even questionable, mainly based on the fact that adenocarcinoma is the most frequent type among nonsmokers. The percentage of cigarettes with filter tips increased from 0.56% in 1950 to over 90% in the 1980s. The nicotine concentration from filter-tip cigarettes is severalfold lower. Hence, the smoker of filter-tip cigarettes tends to inhale more deeply and smoke more intensely, so that the smaller bronchi and bronchio-alveolar regions, the site of origin of adenocarcinomas, are exposed to higher amounts of lung-specific smoke carcinogens (15). Due to the long latency period of lung cancer development, a small proportion of lung cancers investigated today is due to smoking of unfiltered cigarettes in the past decades, when adenocarcinoma may not have been related very strongly to tobacco smoke. The results of the present analysis are in agreement with the epidemiological evidence that in recent years smoke-induced lung tumor formation is not confined to one cell type.

In the past, much information has been collected on the relative abundance of specific types of base pair substitutions in tumors of various organs. For instance, the strong independence of the mutation spectrum of skin

tumors is a result of the unusual abundance of C to T transitions, considered to be a consequence of UV light-induced dimer formation at tandem pyrimidine sites. In colon tumors, 63% of the base pair changes were reported to be G to A transitions; the G to T transversion contributed only 9% (1). In the lung, the ratio was 24%:40%, consistent with DNA adducts from bulky carcinogens such as PAHs, which have been shown to lead predominantly to G to T transversions. In breast tumors, the ratio was 36%:13% (i.e., somewhere in between, indicative of the multifactorial process of carcinogenesis in the mammary gland). Such data, together with the knowledge of the relationship between specific adducts and their mutation spectra, should now be combined with the position-specific information derived from the present analysis in order to take full advantage of the possibilities of the molecular epidemiologic approach.

The focus of the present analysis was the base pair–specific, quantitative, organ-to-organ comparison of the complete p53 mutation spectra. It allowed a number of interesting hypotheses to be formulated. Numerous aspects may have been overlooked, and others may only be discovered with a more refined analysis, especially by stratifying the data by histological type and information on specific exposures. This will be possible as the database increases in size and adopts a more consistent terminology for the entries.

REFERENCES

1. Greenblatt MS, Bennett WP, Hollstein M, Harris CC. Mutations in the p53 tumor suppressor gene: clues to cancer etiology and molecular pathogenesis. Cancer Res 54:4855–4878 (1994).
2. Hollstein M, Shomer B, Greenblatt M, Soussi T, Hovig E, Montesano R, Harris CC. Somatic point mutations in the p53 gene of human tumors and cell lines: updated compilation. Nucleic Acids Res 24:141–146 (1996).
3. Hainaut P, Soussi T, Shomer B, Hollstein M, Greenblatt M, Hovig E, Harris CC, Montesano R. Database of p53 gene somatic mutations in human tumors and cell lines: updated compilation and future prospects. Nucleic Acids Res 25:151–157 (1997).
4. Lasky T, Silbergeld E. p53 Mutations associated with breast, colorectal, liver, lung, and ovarian cancers. Environ Health Perspect 104:1324–1331 (1996).
5. Lutz WK. In vivo covalent binding of organic chemicals to DNA as a quantitative indicator in the process of chemical carcinogenesis. Mutat Res 65:289–356 (1979).
6. Ames B, Gold LS. Endogenous mutagens and the causes of aging and cancer. Mutat Res 250:3–16 (1991).
7. Denissenko MF, Pao A, Tang M, Pfeifer GP. Preferential formation of benzo[a]pyrene adducts at lung cancer mutational hotspots in p53. Science 274:430–432 (1996).
8. Lutz WK, Fekete T. Endogenous and exogenous factors in carcinogenesis: limits to cancer prevention. Int Arch Occup Environ Health 68:120–125 (1996).
9. Dumaz N, Stary A, Soussi T, Daya-Grosjean L, Sarasin A. Can we predict solar ultraviolet radiation as the causal event in human tumours by analysing
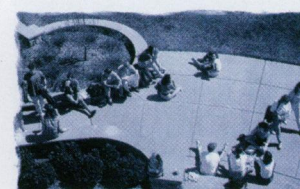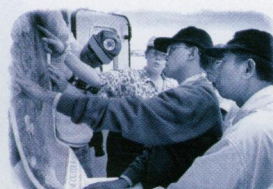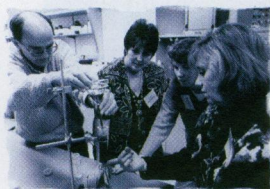
the mutation spectra of the *p53* gene? Mutat Res 307:375–386 (1994).

10. Dayagrosjean L, Dumaz N, Sarasin A. The specificity of *p53* mutation spectra in sunlight induced human cancers. J Photochem Photobiol B Biol 28:115–124 (1995).

11. Doll R. Urban and rural factors in the aetiology of cancer. Int J Cancer 47:803–810 (1991).

12. Wild CP, Kleihues P. Etiology of cancer in humans and animals. Exp Toxicol Pathol 48:95–100 (1996).

13. Doll R, Peto R. The causes of cancer: quantitative estimates of available risks of cancer in the United States today. J Natl Cancer Inst 66:1190–1308 (1981).

14. Shirai T. Etiology of bladder cancer. Semin Urol 11:113–126 (1993).

15. Hoffmann D, Rivenson A, Hecht SS. The biological significance of tobacco-specific N-nitrosamines: smoking and adenocarcinoma of the lung. Crit Rev Toxicol 26:199–211 (1996).

16. Husgafvel-Pursiainen K, Kannio A. Cigarette smoking and p53 mutations in lung cancer and bladder cancer. Environ Health Perspect 104(suppl 3):553–556 (1996).

17. Spruck CH, Rideout WM, Olumi AF, Ohneseit PF, Yang AS, Tsai YC, Nichols PW, Horn T, Hermann GG, Steven K, Ross RK, Yu MC, Jones PA. Distinct pattern of *p53* mutations in bladder cancer: relationship to tobacco usage. Cancer Res 53:1162–1166 (1993).

18. Biggs PJ, Warren W, Venitt S, Stratton MR. Does a genotoxic carcinogen contribute to human breast cancer? The value of mutational spectra in unravelling the aetiology of cancer. Mutagenesis 8:275–283 (1993).

19. Wolff MS, Collman GW, Barrett JC, Huff J. Breast cancer and environmental risk factors: epidemiological and experimental findings. Annu Rev Pharmacol Toxicol 36:573–596 (1996).

20. Wynder EL, Covey LS. Epidemiologic patterns in lung cancer by histologic type. Eur J Cancer Clin Oncol 23:1491–1496 (1987).

21. Devesa SS, Shaw GL, Blot WJ. Changing patterns of lung cancer incidence by histological type. Cancer Epidemiol Biomarkers Prev 1:29–34 (1991).

22. Patel AR, Obrams GI. Adenocarcinoma of the lung. Cancer Epidemiol Biomarkers Prev 4:175–180 (1995).